

ANALYSE DES TRAJECTOIRES DE DÉVELOPPEMENT : VUE D'ENSEMBLE D'UNE MÉTHODE SEMIPARAMÉTRIQUE FONDÉE SUR LE GROUPEMENT

Daniel S. Nagin¹

RÉSUMÉ

Une trajectoire de développement décrit l'évolution d'un comportement au fil des âges ou du temps. La présente communication donnera une vue d'ensemble d'une méthode semi-paramétrique fondée sur le groupement pour analyser les trajectoires de développement. Cette méthode offre une autre solution que celle consistant à supposer que l'on a affaire à une population homogène de trajectoires, comme cela est le cas dans les modèles types de croissance. Quatre capacités seront décrites, à savoir 1) la capacité de reconnaître des groupes distincts de trajectoires plutôt que d'en supposer l'existence, 2) la capacité d'estimer la proportion de la population qui suit chacun de ces groupes de trajectoires, 3) la capacité d'établir le lien entre la probabilité d'appartenir à un groupe ou caractéristiques et les circonstances individuelles et 4) la capacité d'utiliser les probabilités d'appartenir à un groupe à diverses autres fins, telles que la création de profils des membres des groupes. En outre, nous décrirons des extensions importantes de la méthode, c'est-à-dire la capacité d'ajouter des covariables variant en fonction du temps aux modèles de trajectoire, et la capacité d'estimer des modèles à trajectoire collective de comportements distincts, mais associés. La première fournit la capacité statistique de vérifier si un facteur contemporain, tel qu'une intervention expérimentale ou un événement non expérimental comme la grossesse, fait dévier une trajectoire préexistante. La seconde permet d'étudier le déroulement de comportements distincts mais associés, tels que les comportements problématiques durant l'enfance et la toxicomanie à l'adolescence.

1. INTRODUCTION

Le présent article donne un bref aperçu, non technique, de dix années de travaux de recherche en vue d'élaborer une méthode d'analyse des trajectoires de développement fondée sur le groupement. Ces travaux sont décrits de façon plus complète, y compris les détails techniques, dans Nagin (1999) et dans Nagin et Tremblay (2001a). Un exposé complet paraîtra dans un ouvrage qui sera publié par Harvard University Press. En outre, une procédure basée sur SAS comportant les fonctions nécessaires pour soutenir toutes les analyses décrites dans le présent article peut être obtenue gratuitement sur le site Web www.stat.cmu.edu/~bjones/. De la documentation est également disponible sur le site. Pour une vue d'ensemble du logiciel, voir Jones, Nagin et Roeder (2001).

Les données longitudinales constituent le fondement empirique de la recherche portant sur des sujets forts variés en sciences sociales et en science du comportement, ainsi qu'en médecine. Ainsi, les psychologues et les psychiatres utilisent des données longitudinales pour étudier l'évolution des psychopathologies, les criminologues, pour analyser la progression de la délinquance et de la criminalité aux diverses étapes de la vie, les économistes, pour estimer les modèles de la dynamique du revenu, les sociologues, pour étudier l'évolution du statut socio-économique des collectivités et le monde de la recherche médicale, pour étudier l'évolution des maladies et des processus physiologiques.

L'objectif statistique commun de ces divers domaines d'application est la modélisation de l'évolution d'un phénomène en fonction de l'âge ou du temps. Au cours du dernier quart de siècle, deux grandes branches de la méthodologie se sont développées pour l'analyse statistique des processus de croissance, à savoir la modélisation hiérarchique (Bryk et Raudenbush, 1987, 1992; Goldstein, 1995) et l'analyse des courbes latentes (McArdle et Epstein, 1987; Meredith et Tisak, 1990; Muthen, 1989; Willett et Sayer, 1994). Bien que ces deux classes de modèles diffèrent à certains égards importants, elles présentent aussi d'importants points communs (MacCallum, Kim, Malarkey, Kiecolt-Glaser, 1997; Willett et Sayer, 1994; Raudenbush, 2001). Aux fins du présent article, l'un de ces points communs est essentiel : dans l'un et l'autre cas, les distributions non conditionnelle et conditionnelle

¹ Carnegie Mellon University, Pittsburgh, Etats-Unis.

de la population de courbes de croissance sont modélisées à partir de fonctions continues de distribution. Les modèles non conditionnels estiment deux caractéristiques importantes de la distribution de population des paramètres des courbes de croissance, à savoir leur moyenne et leur structure de covariance. La première définit la croissance moyenne dans la population et la seconde cale la variance de la croissance dans la population. Les modèles conditionnels sont conçus pour expliquer cette variabilité en établissant le lien entre les paramètres de croissance et une ou plusieurs variables explicatives.

Le présent article décrit une approche semiparamétrique distincte, fondée sur le groupement, de modélisation des trajectoires de développement. Comme le décrivent Nagin (1999) et Nagin et Tremblay (2001), la méthode repose sur l'élaboration de modèles à mélange fini de distributions. Son application à un ensemble de données recueillies à Montréal est illustrée à la figure 1. Plus précisément, ces données proviennent d'une étude réalisée auprès de 1 037 garçons de race blanche de descendance française. Les sujets ont été sélectionnés en 1984 dans les classes de maternelle des quartiers à faible statut socio-économique de Montréal. Après l'évaluation à l'âge de 6 ans, on a interviewé les garçons chaque année, de l'âge de 10 à 17 ans, ainsi que d'autres personnes bien informées. L'évaluation portait sur un large éventail de facteurs, dont l'autodéclaration par le garçon de son appartenance à une bande de délinquants au cours de l'année ayant précédé l'entrevue.

Telle que décrite dans Lacourse et coll. (à paraître), l'application de la méthode fondée sur le groupement aux données sur l'adhésion à une bande a permis de définir les trois groupes nettement distincts illustrés à la figure 1. La trajectoire de chacun est décrite par la probabilité d'adhésion à une bande à chaque âge : un premier grand groupe, qui, selon les estimations représente 74,4 % de la population, englobe les sujets pour lesquels la probabilité d'adhésion à une bande est très faible à tous les âges. Le deuxième, appelé groupe de la manifestation durant l'enfance, comprend les sujets pour lesquels la probabilité d'adhésion à une bande est forte au départ, à l'âge de 11 ans, augmente modérément jusqu'à l'âge de 14 ans, puis diminue par après. Le troisième, appelé groupe de la manifestation à l'adolescence, comprend les sujets pour lesquels la probabilité d'adhésion à une bande est quasiment nulle à l'âge de 11 ans, mais augmente par après à un taux effectivement supérieur à celui observé pour le groupe de la manifestation durant l'enfance. Selon cette méthode, les deux derniers groupes représentent chacun 13,4 % de la population échantillonnée.

La motivation essentielle de la méthode fondée sur le groupement est illustrée par une gravure sur la modération datant de la fin du XIX^e siècle. La partie gauche de la gravure dépeint un cheminement de vie socialement et personnellement productif. Le premier panneau montre un garçon décidé, se rendant à l'école livres en main. Dans le deuxième panneau, le garçon est devenu un jeune homme et sa femme et son enfant le regardent avec amour pendant qu'il travaille arduement. Dans le troisième panneau, le personnage central est maintenant un vieil homme assis dans un jardin entouré de la famille de son enfant devenu adulte. Le côté droit de la gravure dépeint un cheminement de vie socialement contre-productif et personnellement destructeur. Dans le premier panneau, le garçon boit et fait la fête avec ses amis; dans le deuxième, il bat sa femme dans une rage alcoolique et dans le dernier, il est aux travaux forcés en prison. Abstraction faite du ton moralisateur de cette gravure victorienne, la méthode fondée sur le groupement est conçue pour définir ce genre de trajectoires de vie distinctes et fournir une base statistique pour la découverte des forces qui poussent les individus à les emprunter.

Raudenbush (2001:59) considère la méthode fondée sur le groupement comme une stratégie de modélisation « multinomiale » et offre des éclaircissements précieux au sujet d'une catégorie importante de problèmes pour lesquels elle convient particulièrement bien. Pour certains aspects du développement, il est raisonnable de supposer que les processus de croissance ou de décroissance sont les mêmes pour tous les individus. Il cite comme exemple quintessentiel l'acquisition du langage. Un autre bon exemple est celui de l'évolution du temps passé avec les pairs de l'enfance à l'adolescence (Warr, 2001). Les méthodes classiques des courbes de croissance conviennent bien à l'analyse de ce genre de phénomène de développement. Par contre, il existe de grandes classes de problèmes auxquels les modèles de courbe de croissance ne s'adaptent pas naturellement. Raudenbush choisit la dépression comme exemple. Il fait remarquer qu'il n'est pas sensé de supposer que tout individu suit une trajectoire de dépression croissante (ou décroissante), car nombre de personnes ne seront jamais très déprimées, alors que d'autres le seront sans cesse fortement et d'autres encore deviendront progressivement plus déprimées. Pour cette catégorie de problèmes, il recommande la méthode fondée sur le groupement qui est le sujet du présent article.

L'élément essentiel est celui sur lequel Raudenbush appuie la distinction entre les processus de développement qui sous-tendent l'acquisition du langage et la dépression. Puisque le vocabulaire de tous les jeunes enfants issus d'une

population normale augmente avec l'âge, il est raisonnable de se poser les questions qui suivent. Quelle est la courbe de croissance moyenne du vocabulaire des enfants pendant une tranche d'âge particulière? Quelle est l'importance de la variation entre les courbes individuelles de croissance de l'acquisition du langage? Comment ces variations entre individus sont-elles reliées à des facteurs comme le fonctionnement cognitif de l'enfant et le niveau de scolarité des parents? Comment les variations de l'acquisition du langage chez un individu particulier (« dans un individu ») sont-elles reliées à l'évolution des interactions avec les principaux gardiens à cause, par exemple, de conflits parentaux?

Ces questions sont formulées dans la langue de l'analyse de la variance, comme en témoigne l'utilisation d'expressions telles que « variation dans un individu » et « variation entre individus ». Il est naturel de procéder ainsi, puisque l'analyse classique fondée sur les courbes de croissance a pour origine l'analyse de la variance. Comme cette dernière, l'analyse des courbes de croissance a pour but de trier les facteurs qui expliquent la variation par rapport à une moyenne de population. Elle s'appuie aussi sur l'hypothèse de normalité qui est la pierre angulaire de l'analyse de la variance. Plus précisément, l'analyse des courbes de croissance suppose que les paramètres de croissance au niveau individuel suivent une loi normale dans la population.

Pour concevoir l'analyse statistique d'un processus variable en fonction du temps dans le cadre conceptuel de l'analyse de la variance, il faut avoir la sagesse de caractériser les différences de population concernant un phénomène particulier d'après la variation par rapport à la tendance temporelle de la moyenne. Pour des processus comme l'acquisition du langage, la tendance de la moyenne est, en fait, un point d'ancrage statistique raisonnable pour la description de la variabilité individuelle. Cependant, pour nombre de processus qui évoluent en fonction du temps ou de l'âge, il n'en est pas ainsi. Si nous revenons à l'exemple de la dépression, cela n'a aucun sens de concevoir une analyse statistique des différences de population ayant trait à l'évolution progressive de la dépression d'après la variation par rapport à une « trajectoire de dépression » moyenne. La plupart des formes de psychopathologie, l'usage du tabac, l'alcoolisme et la consommation de drogues illicites sont d'autres exemples de phénomènes comportementaux évolutifs que l'on ne peut décrire correctement d'après la variation par rapport à une moyenne de population. De façon plus générale, l'application d'une méthode fondée sur le groupement pour analyser les mesures répétées est utile pour les phénomènes susceptibles de comprendre des trajectoires de changement en fonction de l'âge ou du temps qui varient qualitativement d'une sous-population à l'autre et qu'on ne peut dégager *ex ante* d'après la mesure de certaines caractéristiques, comme le sexe ou la race.

2. AUTRES RAISONS JUSTIFIANT UN MODÈLE STATISTIQUE FONDÉ SUR LE GROUPEMENT

En psychologie du développement, il existe une longue tradition de théorisation du développement normal et pathologique axée sur les groupes. Mentionnons, à titre d'exemples, les théories du développement de la personnalité (Caspi, 1998), de la toxicomanie (Kandel, 1975), de l'apprentissage (Holyoak et Spellman, 1993), de l'acquisition du langage et de la conceptualisation (Markman, 1989), de la dépression (Kasen et coll., 2001), des troubles de l'alimentation (Tyrka et coll., 2000), de l'anxiété (Cloninger, 1986) et du développement de comportements prosaïques comme la conscience (Kochanska et coll., 1997) et des comportements antisociaux, comme la délinquance (Loeber, 1991; Moffitt, 1993; Patterson, DeBaryshe et Ramsey, 1989).

Puisque les méthodes classiques de modélisation des courbes de croissance sont mal adaptées à leurs besoins, les spécialistes du développement recourent habituellement à l'application de règles d'attribution fondées sur des critères de classification subjectifs pour créer des catégories de trajectoires de développement. Par exemple, Haapasalo et Tremblay (1994) proposent une taxonomie comptant cinq groupes, à savoir les grands bagarreurs stables, les grands bagarreurs qui cessent de se bagarrer, les grands bagarreurs à manifestation tardive, les grands bagarreurs variables et les non-bagarreurs. Ils ont créé ces groupes d'après des évaluations de l'agressivité physique s'étendant de l'âge de 6 ans à l'âge de 12 ans dans le cadre de l'étude longitudinale réalisée à Montréal décrite plus haut. Dans cette étude, lors de chaque évaluation, les chercheurs ont attribué une cote aux sujets au moyen d'une échelle d'agressivité physique comportant les trois éléments suivants : donner des coups de pieds, mordre et frapper d'autres enfants; se battre avec d'autres enfants; et intimider les autres enfants. Pour d'autres renseignements sur l'étude, consulter Tremblay et coll. (1987). Haapasalo et Tremblay ont considéré les garçons obtenant une cote élevée sur cette échelle lors de toute période d'évaluation comme des « grands bagarreurs » pour la période en

question. Puis, ils ont défini des règles de classification des individus dans les cinq groupes taxonomiques susmentionnés. Ces règles étaient fondées sur la fréquence à laquelle chaque garçon avait été considéré comme un grand bagarreur et la tendance à être considéré comme tel. Par exemple, les grands bagarreurs ayant cessé de se bagarrer étaient des garçons qui avaient été de grands bagarreurs à la maternelle, mais qui n'avaient pas été considérés ainsi lors de plus d'une période d'évaluation suivante.

La taxonomie bien connue de Moffitt (1993) est plus parcimonieuse. Elle ne postule que deux trajectoires distinctes de développement du problème comportemental. Un groupe suit ce que l'auteure appelle une trajectoire persistante au cours de la vie (PCL) de comportement antisocial et l'autre, une trajectoire limitée à l'adolescence (LA). Dans les tests empiriques de sa théorie (p. ex., Moffitt, 1996), elle utilise des règles de classification qui sont conceptuellement semblables à celles appliquées par Haapasalo et Tremblay. Les trajectoires PCL sont définies comme correspondant à des individus qui obtiennent une cote supérieure d'au moins un écart-type à la moyenne pour trois des quatre évaluations au moyen d'un indice de trouble des conduites faites entre les âges de 5 et 11 ans et qui obtiennent aussi une cote supérieure d'au moins un écart-type à la moyenne en ce qui concerne l'autodéclaration de la délinquance au moins une fois, à l'âge de 15 ans ou de 18 ans. Les trajectoires LA sont définies comme correspondant à des individus qui ne satisfont pas aux critères PCL pour les troubles des conduites durant l'enfance, mais qui atteignent le seuil PCL de délinquance à l'adolescence.

Bien que ces règles d'affectation soient généralement raisonnables, leur utilisation présente des limites et des écueils. L'un d'eux est qu'il faut supposer a priori qu'il existe des trajectoires de développement distinctes. Donc, l'analyse ne permet pas de vérifier l'existence même de ces trajectoires, ce qui représente un défaut fondamental. Un deuxième écueil, qui est un corollaire du premier, est le risque de « sur et sous-ajuster » simultanément les données en créant des groupes de trajectoires qui reflètent uniquement la variation aléatoire et en omettant de repérer les profils de développement inhabituels, mais néanmoins réels. Troisièmement, la spécification *ex ante* des règles n'offre aucune base pour le calage de la précision des classifications individuelles sur les divers groupes que comprend la taxonomie. Autrement dit, on ne peut quantifier l'incertitude quant à l'appartenance d'un individu à un groupe sous la forme de probabilités et des intervalles de confiance connexes.

Pour illustrer ces limites, considérons l'étude de Haapasalo et Tremblay (1994). La figure 2 donne les résultats d'une application de la méthode fondée sur le groupement au même ensemble de données recueillies à Montréal que celui utilisé dans leur analyse. Plus précisément, la méthode a été appliquée aux évaluations annuelles de l'agressivité physique faites à l'âge de 6 ans, puis de nouveau entre les âges de 10 à 15 ans. Selon cette application, qui a d'abord été décrite dans Nagin et Tremblay (1999), le modèle le mieux ajusté aux données est un modèle à quatre groupes. Un groupe appelé « agressivité faible » comprend les individus qui ne manifestent que peu d'agressivité physique, voire aucune. Il est estimé qu'il représente environ 15 % de la population échantillonnée. Un deuxième groupe, comprenant environ 50 % de la population, est celui de l'« agressivité moyenne à la baisse ». À 6 ans, les garçons de ce groupe manifestaient un niveau moyen d'agressivité physique, mais à 10 ans, ils avaient en grande partie cessé d'être physiquement agressifs. Un troisième groupe, comprenant environ 30 % de la population, est celui de l'« agressivité élevée à la baisse ». Ce groupe comprend des garçons qui ont obtenu une cote élevée d'agressivité physique à l'âge de 6 ans, mais une cote nettement plus faible à 15 ans. Malgré cette nette diminution de l'agressivité, à 15 ans, ils continuaient de manifester un niveau moyen d'agressivité physique. Enfin, il existe un petit groupe d'« agressivité chronique », comprenant environ 5 % de la population, dont les membres ont affiché un haut niveau d'agressivité physique pendant toute la période d'observation.

L'analyse de Nagin et Tremblay offre un fondement statistique formel pour l'existence de trois des groupes hypothésisés dans la taxonomie de Haapasalo et Tremblay, à savoir celui des grands bagarreurs stables qui correspond au groupe à trajectoire d'agressivité chronique, celui des grands bagarreurs ayant cessé qui correspond au groupe ayant une trajectoire à diminution prononcée de l'agressivité et le groupe des non-bagarreurs qui correspond au groupe ayant une trajectoire de faible agressivité. Par contre, il n'existe aucune preuve de l'existence d'une trajectoire correspondant au groupe des grands bagarreurs à manifestation tardive ou d'un groupe de grands bagarreurs variables. Ces dernières classifications causent vraisemblablement un surajustement des données par création de groupes où la variation aléatoire est confondue avec les différences structurelles réelles. Il semble aussi y avoir des preuves d'un problème de sous-ajustement. La taxonomie ne permet pas de repérer le groupe ayant une trajectoire d'agressivité modérée à la baisse. Dans Nagin et Tremblay (1999) et dans Nagin et coll. (à paraître), il est montré que le groupe à trajectoire moyenne à la baisse semble avoir, à certains égards, des caractéristiques distinctes de celles du groupe à trajectoire d'agressivité physique faible.

Les études de Nagin, Farrington et Moffitt (1995) et de Lacourse et coll. (à paraître) sont deux autres exemples de l'utilité de la méthode formelle de groupement des trajectoires comparativement aux méthodes de classification ponctuelles. La première étude visait à tester plusieurs prédictions de la théorie taxonomique de Moffitt. L'un des tests consistait à confirmer l'existence même des trajectoires prévues par la taxonomie. Partant de l'analyse des données sur les condamnations survenues de l'âge de 10 ans à 32 ans chez un échantillon de 403 hommes issus d'un quartier pauvre de Londres, en Angleterre, les auteurs ont dégagé trois trajectoires de délinquance qui sont présentées à la figure 3. L'une de ces trajectoires, qui passe par un sommet aigu à la fin de l'adolescence, concorde étroitement avec le groupe à comportement antisocial limité à l'adolescence prévu par la théorie. La trajectoire en forme de bosse élevée est comparable à certains égards au groupe à comportement antisocial persistant au cours de la vie établi par Moffitt. Le groupe est déjà engagé activement dans la délinquance à l'âge de 10 ans. Cependant, la fréquence des comportements antisociaux, du moins telle que mesurée par les condamnations, dépend fortement de l'âge, un profil que ne prévoit pas la théorie de Moffitt. Elle augmente jusqu'à l'âge d'environ 18 ans, puis commence à diminuer régulièrement. À l'âge de 30 ans, elle est inférieure au niveau de départ observé à l'âge de 10 ans et à peu près égale au niveau observé pour un troisième groupe, appelé délinquants chroniques à faible taux. Ce dernier groupe n'était pas inclus dans la taxonomie de Moffitt. Donc, cette application de la méthode par groupement offre une confirmation élémentaire de l'existence des trajectoires LA et PCL que prévoit Moffitt dans sa taxonomie, mais donne aussi à penser que la trajectoire PCL pourrait dépendre davantage de l'âge que ne laisse supposer la théorie. En outre, la taxonomie de Moffitt ne prévoit pas la trajectoire de comportement antisocial chronique à faible taux.

L'analyse de Lacourse et coll. (à paraître) présentée à la figure 1 illustre deux autres propriétés intéressantes de la méthode de modélisation fondée sur le groupement comparativement à l'application des règles de classification. L'une est la capacité de déceler des évolutions comportementales qualitativement distinctes qu'il n'est pas facile de reconnaître au moyen des règles de classification ponctuelles *ex-ante*. En principe, il est possible de reconnaître *ex-ante* les groupes de manifestation des comportements antisociaux durant l'enfance et à l'adolescence, mais, étant donné leurs trajectoires de développement particulières, il serait fort difficile de les distinguer d'une variation aléatoire sans l'appui d'une méthode statistique formelle. Un deuxième avantage, étroitement lié, émane de l'utilisation d'une structure statistique. C'est grâce à cette structure que la méthodologie permet de faire la distinction entre les variations aléatoires et les différences réelles entre les individus et d'évaluer si la variation à l'intérieur d'un individu est réelle ou uniquement aléatoire. Comme les trajectoires de manifestation durant l'enfance et à l'adolescence sont le produit d'un modèle statistique formel, leur existence repose sur une base plus solide que si elles avaient été construites en s'appuyant sur des règles de classification subjectives.

La méthode fondée sur le groupement permet aussi de répondre aux demandes d'élaboration d'approches « axées sur la personne » pour analyser le développement (Bergman, 1998; Magnusson, 1998). Ces demandes sont motivées par le désir de disposer de méthodes donnant une image statistique instantanée des caractéristiques et comportements distinctifs des individus qui suivent des cheminements de développement différents. La méthode fondée sur le groupement se prête bien à la création de ce genre de profils.

Les statistiques sommaires présentées au tableau 1, qui donnent le profil caractéristique des individus qui suivent les quatre trajectoires d'agressivité physique présentées à la figure 2, illustrent cette capacité. Après avoir estimé les trajectoires, on peut utiliser les estimations des paramètres du modèle pour calculer la probabilité qu'un individu appartienne à chacun des groupes de trajectoires. Pour créer les profils présentés au tableau 1, on a classé les individus dans le groupe de trajectoires auquel ils étaient le plus susceptibles d'appartenir d'après les évaluations chronologiques de leur agressivité physique. Les statistiques sommaires présentées dans le tableau sont simplement le résultat des totalisations croisées de l'adhésion à un groupe ainsi défini et des divers résultats et caractéristiques individuels présentés dans le tableau.

Le profil confirme les observations faites de longue date sur les prédicteurs et les conséquences des problèmes de comportement, comme l'agressivité physique. Les agresseurs chroniques sont les sujets dont les parents ont le niveau de scolarité le plus faible et qui sont le plus fréquemment classés dans le quartile inférieur de la distribution des quotients intellectuels (QI) de l'échantillon. En revanche, les non-agresseurs représentent le groupe le moins susceptible de présenter ces facteurs de risque. En outre, 90 % des agresseurs chroniques n'atteignent pas la huitième année d'études dans les délais prévus et 13 % ont un dossier d'infraction d'adolescent à l'âge de 18 ans. Comparativement, 19 % seulement des non-agresseurs ont pris du retard lorsqu'ils atteignent la huitième année et aucun

ne possède un dossier d'infraction d'adolescent. Entre ces deux groupes, on trouve les agresseurs de niveau faible et les agresseurs de niveau élevé à la baisse.

Le tableau 1 montre que l'appartenance à un groupe de trajectoires varie systématiquement en fonction des caractéristiques de l'individu et des circonstances, comme le niveau de scolarité des parents, qui sont antérieures à la série de mesures. Une généralisation importante du modèle de base permet l'estimation conjointe de la forme des groupes de trajectoires et de l'effet des diverses caractéristiques psychosociales sur la probabilité d'appartenance à un groupe de trajectoires. Cette extension, qui est décrite dans Nagin (1999) et dans Roeder, Lynch et Nagin (1999), permet de déterminer et de tester les prédicteurs précoces des profils de comportement à long terme. Par exemple, une telle analyse montre que la probabilité d'appartenir à un groupe de trajectoires est prédite de façon significative par un faible QI, un faible niveau de scolarité des parents et le fait d'avoir une mère qui a commencé à avoir des enfants à l'adolescence (Nagin et Tremblay, 2001b).

Deux autres extensions importantes du modèle de base méritent d'être mentionnées. L'une, qui est élaborée dans Nagin et coll. (sous presse), permet d'obtenir des *estimations par groupe de trajectoires* de la mesure dans laquelle un événement marquant comme le retard scolaire, une intervention comme le counseling ou un facteur écologique comme l'évolution de la pauvreté du quartier, modifie l'évolution du comportement étudié. La figure 4 illustre cette extension. Elle est le résultat d'une analyse statistique présentée dans Nagin et coll. (sous presse) portant sur l'influence du redoublement d'une année sur l'agressivité physique pour chaque groupe de trajectoires illustré à la figure 2. Un effet statistiquement significatif n'a été observé que pour les deux groupes intermédiaires. La figure 4 donne les résultats du modèle pour le groupe de trajectoires d'agressivité modérée à la baisse. Selon les résultats, l'effet du redoublement d'une année dépendrait de l'âge : le redoublement semble avoir à peu près le même effet aggravant sur l'agressivité physique chez les 6 à 10 ans que chez les 11 à 12 ans, mais ne semble pas avoir d'influence matérielle chez les 13 à 15 ans.

La deuxième extension importante permet d'estimer des modèles à trajectoire conjointe. Deux thèmes importants en psychologie du développement, en psychopathologie du développement et en criminologie du développement sont la comorbidité et la continuité hétérotypique. La comorbidité s'entend de la manifestation contemporaine de deux ou plusieurs états indésirables, comme un trouble des conduites et l'hyperactivité durant l'enfance ou l'anxiété et la dépression à l'âge adulte. La continuité hétérotypique est la manifestation, au cours du temps, d'un trait personnel latent dans des comportements différents, mais analogues.

La comorbidité et la continuité hétérotypique sont habituellement représentées par une statistique sommaire unique, en général un coefficient de corrélation ou un rapport de cotes, qui mesure la co-manifestation des deux comportements ou symptômes étudiés (p. ex., l'hyperactivité et un trouble des conduites à l'âge de 6 ans) ou qui établit le lien entre les deux comportements distincts évalués à diverses étapes de la vie (p. ex., agressivité physique à l'âge de 5 ans et délinquance violente à l'âge de 15 ans). Dans Nagin et Tremblay (2001a), nous élaborons et démontrons une extension du modèle fondé sur le groupement qui établit le lien entre l'évolution longitudinale complète des deux comportements étudiés. La généralisation permet d'estimer conjointement les modèles de trajectoire pour deux séries de mesures distinctes, mais théoriquement reliées. Trois produits importants du modèle sont 1) la forme de la trajectoire des sous-populations distinctes pour les deux séries de mesures, 2) la probabilité d'appartenance à chaque groupe de trajectoires et 3) la probabilité conjointe d'appartenance à des groupes de trajectoires sur l'ensemble des comportements. Ce dernier produit est le progrès important que représente le modèle conjoint, qui offre deux nouvelles fonctions. Premièrement, les probabilités conjointes peuvent caractériser le couplage, au cours du développement, de comportements distincts, mais connexes. Deuxièmement, elles permettent d'évaluer les différences à l'intérieur de la population sous la forme de ce couplage.

Le but de l'article est de donner un aperçu non technique de mon programme d'études de la modélisation des trajectoires fondée sur le groupement. Beaucoup reste à faire. L'élément le plus important est peut-être de permettre de tenir compte du fait que les événements marquants, comme l'échec scolaire, ou les traitements, comme le counseling, peuvent être à la fois une cause et un effet du comportement étudié. Une autre question importante est celle de l'élaboration de lignes directrices rigoureuses permettant de porter des jugements éclairés sur le nombre optimal de groupes qu'il convient de créer. Beaucoup d'efforts seront encore nécessaires afin d'amener la méthode à son plein potentiel, mais je suis maintenant convaincu qu'elle représente un nouvel outil statistique valable pour l'étude des trajectoires de développement.

RÉFÉRENCES

- Bergman, L.R. (1998). A pattern-oriented approach to studying individual development: Snapshots and processes. In R.B. Cairns, L.R. Bergman, and J. Kagan (eds.), Methods and Models for Studying the Individual (pp. 83-122). Thousand Oaks: Sage Publications.
- Bryk, A.S., et Raudenbush, S.W. (1987). Application of hierarchical linear models to assessing change. *Psychology Bulletin*, 101:147-158.
- Bryk, A.S., et Raudenbush, S.W. (1992). Hierarchical Linear Models for Social and Behavioral Research: Application and Data Analysis Methods. Newbury Park, CA: Sage.
- Caspi, A. (1998). Personality development across the life course. In Handbook of Child Psychology, Fifth Edition, ed. N. Eisenberg, (W. Daom, Series Ed.). Volume 3: Social, Emotional, and Personality Development (pp. 311-388). New York: Wiley.
- Cloninger, R.C. (1986). A unified biosocial theory of personality and its role in the development of anxiety states. *Psychiatric Developments*, 3:167-225.
- Cloninger, R.C. (1987). A systematic method for clinical description and classification of personality variants. *Archives of General Psychiatry*, 44:573-588.
- Goldstein, H. (1995). Multilevel Statistical Models, second edition. London: Edward Arnold.
- Haapasalo, J., et Tremblay, R.E. (1994). Physically aggressive boys from ages 6 to 12: Family background, parenting behavior, and prediction of delinquency. *Journal of Consulting and Clinical Psychology*, 62(5), 1044-1052.
- Heckman, J. et Robb, R. (1985). Alternative methods for evaluating the impact of interventions. In Longitudinal Analysis of Labor Market Data, edited by J. Heckman and B. Singer, New York: Cambridge University Press, pp.156-245.
- Jones, B.L., Nagin, D. et Roeder, K. (2001), "A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories," *Sociological Research and Methods*, 2001, 29: 374-393.
- Kandel, D.B. (1975). Stages in adolescent involvement in drug use. *Science*, 190:912-914.
- Kasen S., Cohen, P., Skodol, A.E., Johnson, J.G., Smailes, E. et Brook, J.S. (2001). Childhood depression and adult personality disorder - Alternative pathways of continuity. *Archives of General Psychiatry*, 58:231-236.
- Kochanska, G., Murray, K. et Coy, K.C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development*, 68: 263-277.
- Lacourse, E. Nagin, D., Vitaro, F., Claes, M. et Tremblay, R.E. " Developmental Trajectories of Boys Delinquent Group Membership and Facilitation of Violent Behaviors during Adolescence," , *Development and Psychopathology*, forthcoming.
- _____ (1991). Questions and advances in the study of developmental pathways. In D Cicchetti and S. Toth (eds.), Models and Integrations. Rochester Symposium on Developmental Psychopathology, (Vol. 3, pp. 97-115). Rochester: University of Rochester Press.
- MacCallum, R.C., Kim, C., Malarkey, W.B., et Kiecolt-Glaser, J.K. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research*, 32: 215-253.
- Maddala, G.S. (1983). Limited-Dependent and Qualitative Variables in Econometrics. New York: Cambridge University Press.

- Magnusson, D. (1998). The logic and implications of a person-oriented approach. In R.B. Cairns, L.R. Bergman and J. Kagan (eds.), Methods and Models for Studying the Individual. (pp. 33-64). Thousand Oaks: Sage Publications.
- McArdle, J. J. et Epstein, D. (1997), Latent growth curves within developmental structural equation models. *Child Development*, 58, 110-133.
- Markman, E.M. (1989). Categorization and Naming in Children: Problems of Induction. Cambridge, MA: MIT Press.
- Meredith, W. et Tisak, J. (1990). Latent curve analysis. *Psychometrika*, 55(1):107-122.
- Moffitt, T.E. (1993). Adolescence-limited and life-course persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100(4):674-701.
- Muthen, B.O. (1989). Latent variable modeling in heterogeneous populations. Psychometrika, 54(4):557-585.
- Nagin, D. (1999). Analyzing Developmental Trajectories: Semi-parametric, Group-based Approach. *Psychological Method*, 4,139-177.
- Nagin, D., Farrington, D. et Moffitt, T. (1995). Life-course trajectories of different types of offenders. *Criminology*, 33:111-139.
- Nagin, D.S. et Land, K.C. (1993). Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson model. *Criminology*, 31:327-362.
- Nagin, D., Pagani, L., Tremblay, R. et Vitaro, F. (forthcoming). "Life course turning points: A case study of the effect of grade retention on interpersonal violence. *Development and Psychopathology*.
- Nagin, D.S. et Tremblay, R.E.. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development*, 70:1181-1196.
- _____ (2001a). "Analyzing Developmental Trajectories of Distinct but Related Behaviors: A Group-Based Method," Psychological Methods, 6: 18:34.
- _____ (2001b). "Parental and Early Childhood Predictors of Persistent Physical Aggression in Boys from Kindergarten to High School" , Archives of General Psychiatry, 58:389-394.
- Patterson, G.R., DeBaryshe, B.D. et Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist*, 44:329-335.
- Raudenbush, S.W. (2001). Comparing-personal trajectories and drawing causal inferences from longitudinal data, *Annual Review of Psychology*, 52: 501-25.
- Roeder, K., Lynch, K. et Nagin, D. (1999). Modeling uncertainty in latent class membership: A case study in criminology," *Journal of the American Statistical Association*, 94:766-776.
- Tremblay, R.E., Desmarais-Gervais, L., Gagnon, C. et Charlebois, P. (1987). The preschool behavior questionnaire: stability of its factor structure between culture, sexes, ages and socioeconomic classes. International Journal of Behavioral Development, 10, 467-484.
- Willett, J.B. et Sayer, A.G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, 116(2):363-381.

Tableau 1 : Profils des groupes d'agressivité physique

Variable	Groupe			
	Pas d'agressivité	Agressivité faible à la baisse	Agressivité forte à la baisse	Agressivité chronique
Nombre d'années d'études - Mère	11,1	10,8	9,8	8,4
Nombre d'années d'études - Père	11,5	10,7	9,8	9,1
QI faible (%)	21,6	26,8	44,5	46,4
A terminé la 8 ^e année dans les délais normaux (%)	80,3	64,6	31,8	6,5
Dossier d'infraction d'adolescent (%)	0,0	2,0	6,0	13,3
N ^{bre} de partenaires sexuels à l'âge de 17 ans (année précédente)	1,2	1,7	2,2	3,5

Figure 1: Trajectoires de l'adhésion à une bande

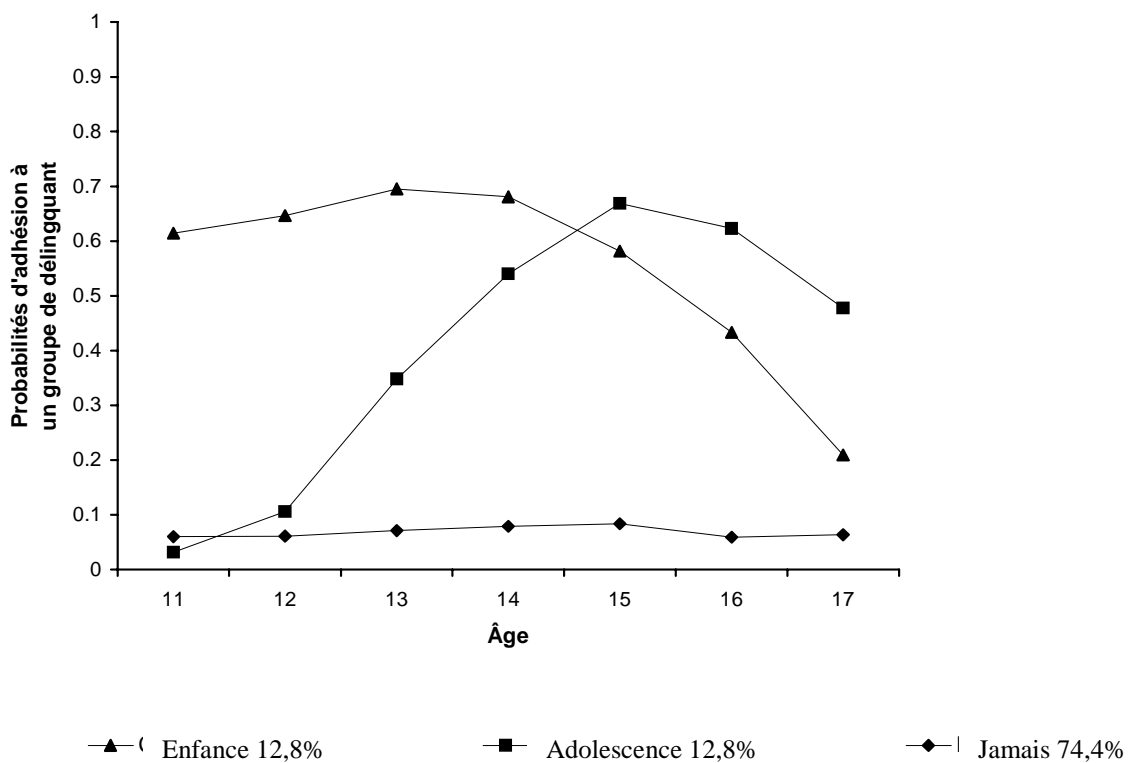


Figure 2 : Trajectoires de l'agressivité physique

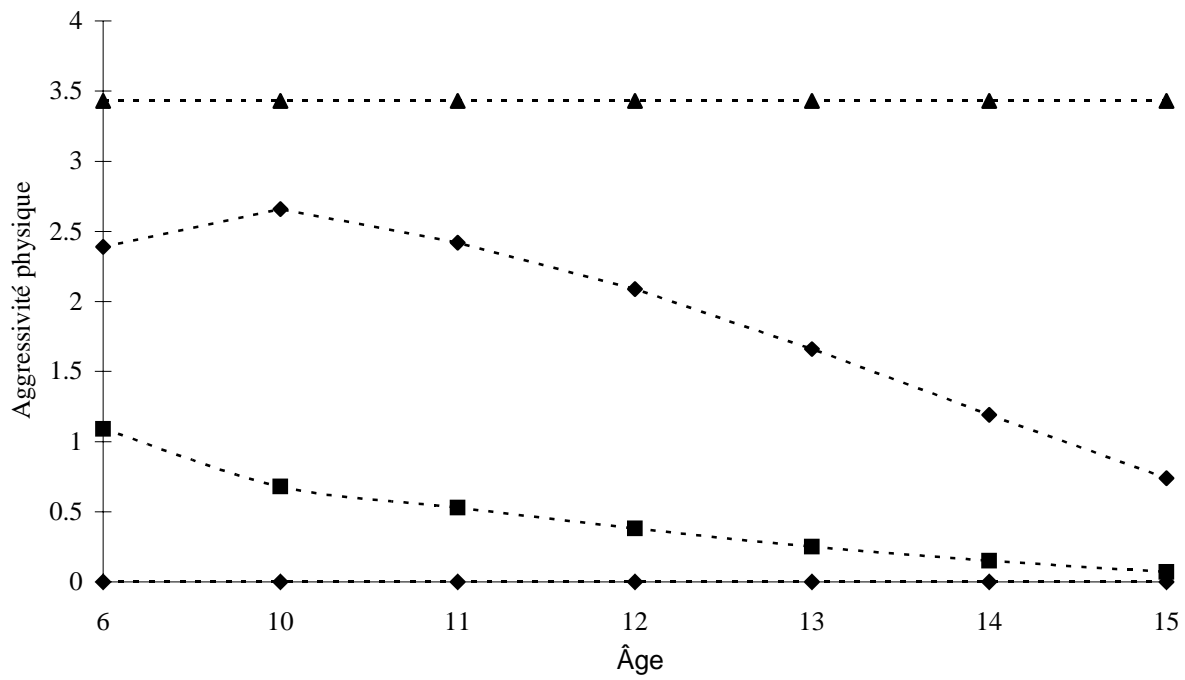


Figure 3 : Trajectoires des condamnations

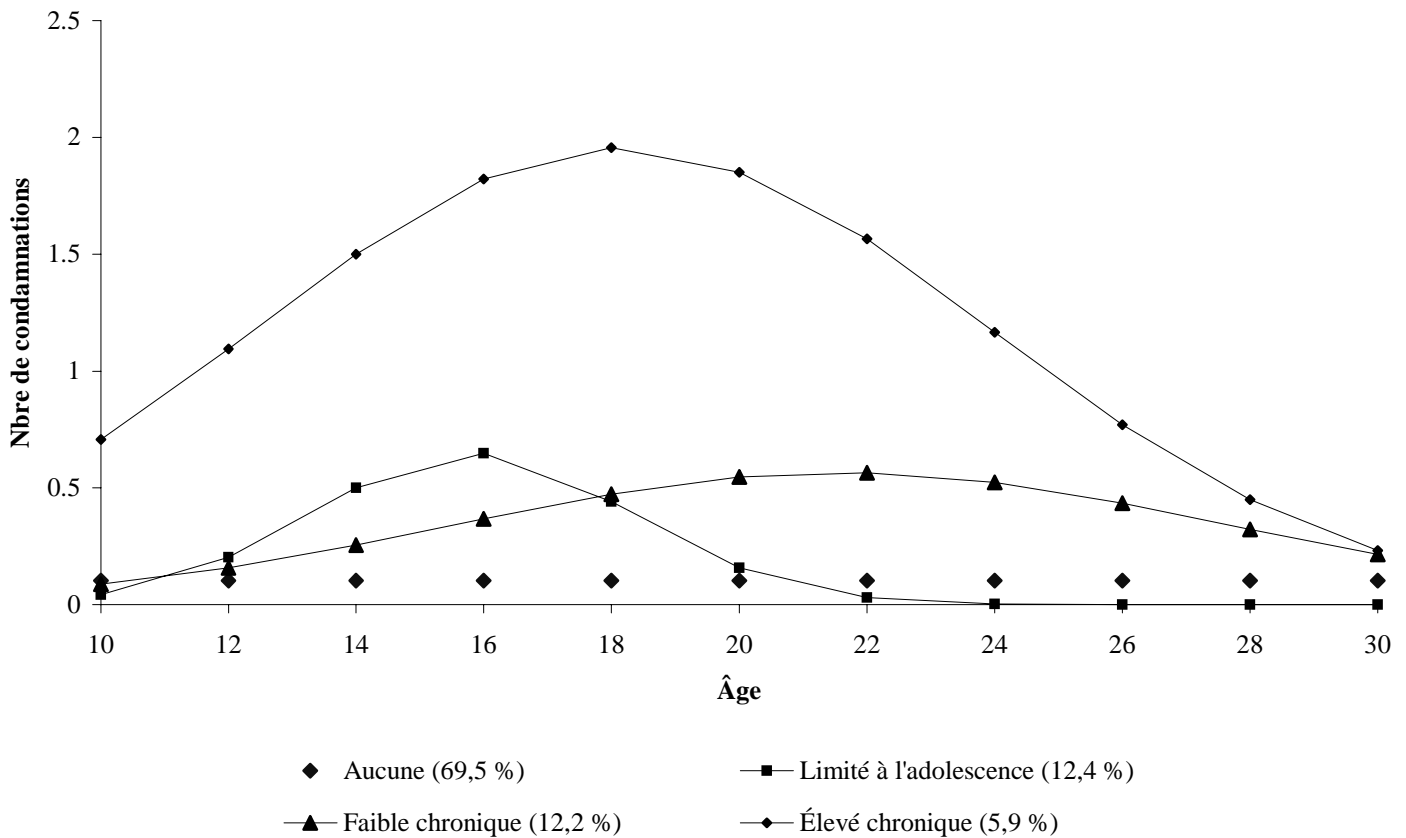


Figure 4: Effets du redoublement scolaire: trajectoire d'agressivité moyenne à la baisse

